REMARKS

Claims 1-44 are pending in the application. Claims 1, 12, 17, and 23 have been amended. No claims have been allowed. Applicants respectfully request entry of the foregoing amendments and allowance of the application as amended.

Petition for extension of time

A petition for a three month extension of time under 37 C.F.R. § 1.136(a) is included herewith, as well as the fee under 37 C.F.R. § 1.117(a)(3).

Double Patenting Rejection

Claims 1-44 were rejected under the judicially created doctrine of obviousness-type double patenting as being unpatentable over claims 1-39 of U.S. Patent application publication number 2003/0228023 (U.S. Patent Application No. 10/400,282). Applicants submit herewith a Terminal Disclaimer in compliance with 37 C.F.R. § 1.130(c), as well as the fee under C.F.R. § 1.20(d). Applicants respectfully request withdrawal of the rejection.

Rejections under 35 U.S.C. § 102

Claims 1-3, 5-18, 20-30, and 33-34 were rejected under 35 U.S.C. § 102(b) as being anticipated by Holzrichter (U.S. patent number 5,729,694, hereinafter "Holzrichter"). Applicants respectfully traverse the rejection.

The examiner states in paragraph 5 of the Office action:

5. Claim1-3, 5-18, 20-30 and 33-44 are rejected under 35 U.S.C. 102(b) as being anticipated by Holzrichter (US PAT. 5,729,694).

Consider claim 1 Holzrihter teaches a method for removing noise from acoustic signals, comprising:

receiving (see fig.43, 52) a plurality of acoustic signals;

receiving (43) information on the vibration of human tissue associated with human voicing activity;

generating at least one first transfer function (57) representative of the plurality of acoustic signals upon determining that voicing information is absent (such as, unvoice) from the plurality of acoustic signals for at least one specified period of time (such as, time frames)(see col.28 line 38-48); and

removing noise (removing noise is inherent to speech recognition algorithm to extract the best speech feature and avoid noise) from the plurality of acoustic signals using the first transfer (57) function to produce at least one denoised acoustic data stream (60, see col. 15 line 29-col. 16 line 3 and col. 60 line 19-30).

Applicants respectfully submit that claims 1-3, 5-18, 20-30, and 33-34 are not anticipated by Holzrichter. Holzrichter lacks at least one element of the claims. Specifically, Holzrichter lacks at least the following elements: transfer function(s); and use of more than one microphone. To explain the differences between the claimed invention and Holzrichter, and to support Applicants statement that Holzrichter lacks the stated elements, Applicants submit the following ananlysis.

Applicants submit that while Holzrichter does discuss receiving both acoustic (microphone-based) and "EM wave" (col. 15, lines 19) based measurements, Holzrichter uses only a single microphone. Further, Holzrichter uses the "EM wave" only to measure "the conditions of the vocal folds and the glottal tissue surrounding the vocal fold structure" (col. 15, lines 26-28). Applicants do not claim to measure the conditions of vocal folds or surrounding structure; but rather claim the use of physiologically-based device to determine the VAD signal, which is not mentioned in Holzrichter. In addition, in contrast to Holzrichter , Applicants claim "receiving a plurality of acoustic signals, wherein receiving the plurality of acoustic signals

includes receiving using a plurality of independently located microphones." (Claim 1 as amended).

Holzrichter does not generate any transfer function as in (57). Rather, Holzrichter generates a Fourier Transform, which is a measure of the frequency content of a signal. In order to generate a transfer function, two signals (an input and output) are needed. Holzrichter has only one signal and simply transforms it into the frequency domain, which is completely different from, and exclusive of, taking two signals and generating a transfer function. Even the quoted passage at col. 28 lines 38-48 does not discuss time frames – it is simply a discussion of how Holzrichter thinks the end of a speech period should be calculated when a voiced to unvoiced transition takes place:

> **In the case that the speech changes from voiced to unvoiced, the last glottal open/close period of the voiced speech sequence has no "next" glottal cycle to use to define its end of period. In one approach, the algorithm continually tests the length of each glottal closed-time in each time frame for excessive length (e.g. 20% longer than the preceding glottal period closure-time). If the period is texted to be too long, the algorithm terminates the period and assigns, for example, a glottal-closure time-duration equal to the fractional closure time of the glottal function measured in the preceding time frames.**

The foregoing is not related to segmenting data into time frames for processing.

Finally, (57) and (60) along with col. 15 line 29 – col. 16 line 3 and col. 60 lines 19-30 have nothing to do with removing noise. (57) is a Fourier Transform, as noted above, and (60) is a "Vocal tract feature extractor" which is unrelated to present claims. Col. 15 lines 29-col. 16 line 3 describes Holzrichter's perception of how a "vocal tract Fourier Transform" would appear (col. 15 line 67-col. 16 line 1). This is to be distinguished as being completely different from a calculating transfer function as claimed.

FIG. 5 shows a system in which knowledge of the vocalized excitation function is used to deconvolve the speech vocal tract transfer function information from measured acoustic speech output each time frame. All of the information gathered during each speech time frame, including: acoustics, EM sensor information, and deconvolved transfer function information, can be processed, normalized, quantized, and stored (along with control information) in a feature vector representing the speaker's voice during one or more speech time frames. Similar deconvolving procedures are used with unvoiced excitation functions. As shown in FIG. 5, an EM sensor control unit 40 drives a repetition rate trigger 41, which drives pulse generator 42, which transmits one or more pulses from antenna 43. EM sensor control unit 40 sets the pulse format, time frame interval, integration times, memory locations, function forms, and controls and initializes pulse generator 42. Control unit 40 and trigger 41 also actuate switch 45 through delay 44 to range gate received pulses. Antenna 43 is positioned to direct transmitted pulses towards the vocal organs and receive pulses reflected therefrom. The received pulses pass through switch 45 and are integrated by integrator 46, then amplified by amplifier 47, and passed through a high pass filter 48 to a processing unit 49. Processing unit 49 contains an AD converter for digitizing the EM signals and also includes zero location detector, memory detector, and obtains glottal area versus time. The digitized and processed data from unit 49 is stored in memory bins 50, from which excitation function feature vectors are formed in block 51. Simultaneously, signals from an acoustic microphone 52 are digitized by AD converter 53, which is also controlled and synchronized by EM sensor control unit 40. The digitized data from AD converter 53 is stored in memory bins 54 from which acoustic feature vectors are formed in block 55. The digitized vocal fold data from memory bins 50 is used to produce a glottal Fourier transform 56, while the digitized acoustic data in memory bin 54 is used to produce an acoustic Fourier transform 57. The two Fourier transforms 56, 57 are deconvolved in block 58 to produce a vocal tract Fourier transform 59 which is then fit to a prechosen functional form to form a vocal tract feature vector in block 60.

Finally, in col. 60 lines 19-30, Holzrichter discusses removing "acoustically generated noise" from the "glottal signal" using "Fourier transform techniques" (col. 60 lines 19-22). This is not related to the present claims because:

1. The noisy signal of Holzrichter is a "glottal signal", while the claimed noisy signal is an acoustic microphone signal.
2. Holzrichter does not describe or embody his "Fourier transform techniques", so they are not available to be compared to the claimed techniques.

> Acoustically generated noise can be removed from the glottal signal by using microphone information to subtract the noise signal, or by using Fourier transform techniques to filter out acoustic signals from the glottal motion signals.
>
> The functional shape of the volume air flow excitation function in real time, and in transform space (Fourier or Z transform), can be approximated, including the glottal zero (or closed) time. An excitation feature vector is constructed by defining an approximation functional (or table) to the measured excitation function and by obtaining a series of numerical coefficients that describe the functional fitting to the numerical data for the defined time frame(s).

The last part (col. 60 lines 23-30) discusses how to approximate a "volume flow excitation function" using "functional [sic] (or table)". This has no bearing on the claims. For example, the claims do not include a "volume flow excitation function". In addition, Applicants respectfully submit that Holzrichter's "functional [sic] (or table)" is described with enough specificity to determine what is being referred to.

The examiner also states:

> Consider claims 2-3 Holzrihter teaches the method of removing noise further comprises:
>
> generating at least one second transfer function (see fig.5 (56)) representative of the plurality of acoustic signals upon determining that voicing information is present in the plurality of acoustic signals for the at least one specified period of time (such as, time frames)(see col.28 line 38-48), and removing noise (removing noise is inherent to speech recognition algorithm to extract the best speech feature and avoid noise) from the plurality of acoustic signals using at least one combination (58) of the at least one first transfer function (57) and the at least one second transfer function (56) to produce at least one denoised acoustic data stream(see col. 15 line 29-col. 16 line 3 and col. 60 line 19-30); and the plurality of acoustic signals include inherently (because the EM sensor 43 and microphone 52 picks up the noise source signal and the acoustic signal) at least one reflection of at least one associated noise source signal and at least one reflection of at least one acoustic source signal (see col. 14 line 46-67 and col.24 line 29-61).

Applicant respectfully submits that Holzrichter only mentions noise removal in a single context. Specifically, Holzrichter desires to remove "acoustically generated noise" from the "glottal signal" using "Fourier transform techniques" (col. 60 lines 19-22). Not once, in any context, does Holzrichter mention the removal of acoustic noise from a microphone signal using a second microphone signal. The references to (col. 28 lines 38-48) and (col. 15 lines 29-col. 16 line 3) and (col. 60 lines 19-30) have been shown to be unrelated above. Col. 14 lines 46-67 states:

> FIGS. 3A and FIG. 3B show two types of laboratory apparatus for measuring the simultaneous properties of several speech organs using EM sensors and for obtaining simultaneous acoustic information. FIG. 3A, in particular, shows highly accurate laboratory instrumentation assembled to obtain very high fidelity, linear, and very large dynamic range information on the vocal system during each speech time frame. FIG. 3A shows a view of a head with three antennas **21, 22, 23** and an acoustic microphone **24** mounted on a support stand **25**. Antennas **21, 22, 23** are connected to pulse generators **26***a, b, c* through transmit/receiver switches **27***a, b, c* respectively. Pulse generators **26***a, b, c* apply pulses to antennas **21, 22, 23**, which are directed to various parts of the vocal system. Antennas **21, 22, 23** pick up reflected pulses, which are then transmitted back through switches **27***a, b, c* to pulse receivers and digitizers (e.g., sample and hold units) **28***a, b, c*. Acoustic information from microphone **24** is also input into pulse receiver and digitizer **28d**. Support stand **25** positions the antennas **21, 22, 23** to detect signals from various parts of the vocal tract, e.g., by using face positioning structure **29** and chest positioning structure **30**. As shown, antenna **21** is positioned to detect the tongue, lip,

This is a simple description of the multiple antenna and SINGLE microphone configuration envisioned by Holzrichter. In contrast, Applicants claim the use of multiple microphones.

Col. 24 lines 29-61 state:

> 3) Remove post-glottal pressure induced vibrations of glottal tissue and nearby tissue from the EM sensor signal, and therewith from the associated model of volume air flow versus sensor signal. Use one of two related methods. Method 3A) Filter the raw EM sensor excitation signal using transform or circuit techniques to remove the acoustic pressure induced higher frequency noise, but preserve the needed low frequency excitation function shape information for model generated values of volume air flow and for subsequent feature vector formation. Method 3B) Use the tissue vibration signal from the EM sensor and the acoustic output (corrected for timing delays) to determine the backward acoustic transfer function. Divide the Fourier transforms of the vibration signal by that of the acoustic signal, and store the numerical (or curve fit) transfer function information in memory for recall as needed. Next, for each time frame, use the backward transfer function to calculate the glottal tissue vibration level associated with the measured output acoustic signal. Then subtract the backward transferred acoustic signal from the EM-sensor generated and processed signal, to obtain a "noise free" excitation function signal. This signal represents a backward traveling acoustic sound wave that induces mechanical vibrations of glottal tissue and nearby air tract tissues in directions transverse to the air flow. This acoustic wave has little effect on the positions of the vocal fold edges, and thus it does not affect the actual volume air flow, U. However, certain EM sensors do measure this noise, and it shows up on the EM signal describing the excitation function (see FIG. 4B for an example). This noise level is found to be speaker specific. For high fidelity, speaker independent excitation function coding, such vibration signals mixed with the gross air flow values are undesirable.

In this passage, Holzrichter discusses how he envisions removing "post-glottal pressure induced vibrations" (i.e. the speech of the user) from the "EM sensor signal" (col. 24 lines 29-30). Holzrichter lack elements of the claimed invention, and even teaches away from the claimed invention. For example:

1. The claimed signal of interest is the user's speech, whereas the user's speech is "noise" in Holzrichter.
2. Holzrichter's signal of interest is the "EM sensor signal" which Holzrichter desires to associate with volume air flow.

3. The claimed noise is environmental acoustic noise, whereas the speech of the user is noise in Holzrichter.

The Office action continues:

> Consider claims 5-8 Holzrichter teaches that the method of removing noise further includes generating at least one third transfer function (see fig.5 (59)) using the at least one first transfer function (57) and the at least one second transfer function (56); the method of generating the at least one first transfer function (see fig.5 (57)) comprises recalculating the at least one first transfer function during at least one prespecified interval (see col. 19 line 26-col. 20 line 15); and the method of generating the at least one second transfer function (see fig.5 (56)) comprises recalculating the at least one second transfer function during at least one prespecified interval (see col. 19 line 26-col. 20 line 15); and the method of generating the at least one first transfer function (see fig.5 (57) comprises use of at least one technique selected from a group consisting of adaptive techniques and recursive techniques (see col. 19 line 26-col. 20 line 15).

As shown above, (56, 57, and 59) are Fourier transforms, not transfer functions as in the claims. Col. 19 line 26 – col. 20 line 15 states:

The feature vector shown in FIG. 12A for the sound /ah/, was constructed using a total of p feature vector coefficients, c_1 through c_p, to describe the processed data. In this example, c_1 is used to describe the type of transfer functions used, e.g. "1" means the use of an ARMA functional in the "pole" and "zero" formulation; c_2 describes the number of "poles" and c_3 describes the number of "zeros" used for the fitting; c_4 indicates the kind of speech unit being spoken, e.g. "0" means isolated phoneme; c_5 describes the type of connection to a preceding acoustic sound unit to be used, e.g. "0" means a connection to the silence phoneme is needed; c_6 describes the connection to the following unit, e.g. "0" means a connection to a following silence phoneme is needed; c_7 describes the 300 ms multi-frame speech segment envelope; c_8 is the pitch (e.g., 120 vocal fold cycles/sec.); and c_9 describes the bandwidth of the fundamental harmonic. Other feature vector coefficients that describe the relative ratios of the 2nd through the 10th harmonic power to the first harmonic, are taken from the power transform of the vocal excitation (FIG. 10B). In addition the fall of the harmonic excitation power per octave, above 1 kHz, can be described by a line with −12 db/octave negative slope. The "pole" and "zero" coefficient data (FIG. 12B) are shown and stored as appropriate coefficients in the vector in FIG. 12A. The last coefficient c_p is the symbol for the sound, and the next to last c_{p-1} is acoustic information from a CASR or similar system which is the acoustic energy per frame. If the user desires to use the alternative formulation of the ARMA transfer functional, the "a" and "b" coefficients can be used (see FIG. 12C).

An alternative approach to describe the feature vector for the "long" speech segment /ah/ is to perform Fourier transformations each 8.3 ms (the period for 120 Hz excitation), and to join 36 individual pitch period frames into a 300 ms long multiple frame speech segment. A second alternative approach would be to take the Fourier transform of the entire 300 ms segment, since it was tested to be constant; however the FFT algorithm would need to handle the large amount of data. Because of the constancy of the acoustic phoneme unit /ah/, the user chose to define the 300 ms period of constancy first, and to then process (i.e., FFT) the repetitive excitation

> and output acoustic signal with a convenient 10 ms period 30 times, and then average the results.
>
> As a test (see Section below on Speech Synthesis) a synthetic speech segment was reconstructed from information in a vector like the one shown in FIG. 12A. The vocal fold excitation function was first reconstructed using the harmonic amplitude and phase information to generate a source term over an interval of 100 ms. The excitation function was sampled at 11 kHz or higher. The time sampled sequence was used to drive the ARMA model specified by a difference equation with poles and zeros. The output of the ARMA model was used to reconstruct the speech sound /ah/ as shown in the section on Speech Synthesis (see FIG. 19), and a pleasing sound, /ah/, was generated and heard by the user.

Again as before, Holzrichter envisions calculating a "feature vector" to describe the speech excitation (using Fourier transforms and other unnamed techniques), and the vocal tract transfer function (a model of the configuration of the vocal tract), and the speech itself. None of this is relevant to the claims, which do not calculate, or approximate, or model the speech excitation or vocal tract transfer function in any way. Holzrichter simply lacks any teaching or suggestion regarding the methods, including transfer functions, as claimed.

Continuing:

> Consider claims 9-11 Holzrichter teaches that the method of information on the vibration of human tissue is provided by a mechanical sensor (such as, motion sensor) in contact with the skin (see figs 3a-3b(29,30,33)) and see col. 14 line 46-col. 15 line 18); and the method of information on the vibration of human tissue is provided via at least one sensor selected from among at least one of an accelerometer, a skin surface microphone in physical contact with skin of a user, a human tissue vibration detector, a radio frequency (R.F) vibration detector, and a laser vibration detector(see figs 3a-3b(29,30,33) and see col. 14 line 46-col. 15 line 18); and the human tissue is at least one of on a surface of a head, near the surface of the head, on a surface of a neck, near the surface of the neck, on a surface of a chest, and near the surface of the chest(see figs 3a-3b(29,30,33)) and see col. 14 line 46-col. 15 line 18).

The quoted passage (col. 14 line 46 – col. 15 line 18) does not, in fact, describe any mechanical sensor in contact with the skin. Nor does the quoted passage, according to Holzrichter, detect vibrations associated with speech. According to Holzrichter (col. 14 lines 47-48):

> **apparatus for measuring the simultaneous properties of several speech organs using EM sensors and for obtaining**

whereby the speech organs are listed as the (col. 14 line 67 - col. 15 line 3),

> **As shown, antenna 21 is positioned to detect the tongue, lip, velum, etc. Antenna 22 is positioned to detect tongue and jaw motion and antenna 23 is position to detect vocal fold motion.**

and "properties" is not explained. Holzrichter simply lacks any teaching or suggestion regarding the use of mechanical or EM sensors to detect skin vibration due to user speech as is in the claims.

Continuing:

> Consider claim 12 Holzrichter teaches that a method for removing noise from electronic signals, comprising:
>
> detecting (see fig.5, (43, EM sensor)) an absence (unvoice) of voiced information during at least one period (see col. 28 line 38-48), wherein detecting includes measuring the vibration of human tissue (see col. 5 line 66-col. 6 line 55); receiving at least one noise source signal during the at least one period (see col. 24 line 29-61); generating at least one transfer function (57) representative of the at least one noise source signal; receiving at least one composite signal comprising acoustic and noise signals; and removing the noise (removing noise is inherent to speech recognition algorithm to extract the best speech feature and avoid noise) signal from the at least one composite signal using the at least one transfer function to produce at least one denoised acoustic data stream(60, see col. 15 line 29-col. 16 line 3 and col. 60 line 19-30).

As shown above, Holzrichter never mentions generating a transfer function – (57) is a Fourier transform, a completely different process. Also, as discussed above, Holzrichter's signal of interest and "noise" are completely different from the claimed

signal of interest and the claimed noise. Applicants have also shown that Holzrichter does not teach detecting vibrations associated with the user's speech.

Continuing:

> **Consider claims 18 and 20 Holzrichter teaches that the method of receiving includes receiving the at least one noise source signal using at least one microphone (see fig.5 (52)); and the method of removing the noise signal from the at least one composite signal using the at least one transfer function (see fig.5 (59)) includes generating at least one other transfer function (57) using the at least one transfer function (see col. 15 line 29-col. 16 line 3).**

Again, Holzrichter uses a single microphone, and never considers that the microphone signal might be polluted with environmental noise. In contrast, environmental noise is at the center of the present claims. Holzrichter never generates a transfer function, only using Fourier transforms, and the only "noise" discussed by Holzrichter is the speech of the user. Applicants emphasize that the speech of the user is what Applicants invention is intended to keep. Further, Holzrichter does not detail the manner in which he would remove the "noise" in any way.

Consider claim 23 Hotzrichter teaches a method for removing noise from electronic signals, comprising:

determining (see fig.5 (40)) at least one unvoicing period during which voiced information is absent (such as, unvoice) based on vibration of human tissue;

receiving (43, 52) at least one noise signal input during the at least one unvoicing period (see col.28 line 38-48) and generating at least one unvoicing transfer function(56) representative of the at least one noise signal (see col. 24 line 29-61);

receiving (43,52) at least one composite signal comprising acoustic and noise signals; and removing the noise signal (removing noise is inherent to speech recognition algorithm to extract the best speech feature and avoid noise) from the at least one composite signal using the at least unvoicing transfer function to produce at least one denoised acoustic data stream (60, see col. 15 line 29-col. 16 line 3 and col 60 line 19-30).

Holzrichter fails to teach or suggest removing acoustic environmental noise from the user's speech, which is a subject matter of the claimed invention. Figure 5 (40) is the "EM sensor control unit", which is not enumerated and which has nothing to do with determining unvoiced periods. Figure 5 (43) and (52) do not receive noise; rather they receive "properties of several speech organs" (col. 14 lines 47-48) and the (not-noisy) speech of the user. As shown above, the only "noise" in Holzrichter's system is the user's speech, which is the claimed signal of interest. Holzrichter further lacks any teaching regarding transfer functions, but rather teaches only the use of Fourier transforms.

> Consider claim 24 Holzrichter teaches that the method of producing at least one denoised acoustic data stream further includes:
>
> determining (see fig.5 (40)) at least one voicing period during which voiced information is present; receiving (52) at least one acoustic signal input from at least one signal sensing device during the at least one voicing period (see col.28 line 38-48) and generating at least one voicing transfer function (57) representative of the at least one acoustic signal; and removing the noise signal from the at least one composite signal using at least one combination of the at least one unvoicing transfer function (56) and the at least one voicing transfer function (57) to produce the denoised acoustic data stream (60, see col. 15 line 29-col. 16 line 3 and col. 60 line 19-30).

Holzrichter never discusses producing a denoised acoustic data stream as claimed. Instead, Holzrichter discusses "removing 'post-glottal pressure induced vibrations' (i.e. the speech of the user) from the 'EM sensor signal' " (col. 24 lines 29-30). Thus, Holzrichter not only lacks elements of the claims, but is not even directed toward achieving a result similar to that achieved by the claimed invention.

Consider claim 26 Holzrichter teaches a system for removing noise from the acoustic signals, comprising:

at least one receiver (see fig.5 (52)) that receives at least one acoustic signal;

at least one sensor (43) that receives human tissue vibration information associated with human voicing activity;

at least one processor (see fig.3b (proceeding electronics)) coupled among the at least one receiver and the at least one sensor (52,43) that generates a plurality of transfer functions (56, 57,59), wherein at least one first transfer function (57) representative of the at least one acoustic signal is generated in response to a determination that voicing information is absent (unvoice) from the at least one acoustic signal for at least one specified period of time (such as, time frames)(see col.28 line 38-48), wherein noise is removed (removing noise is inherent to speech recognition algorithm to extract the best speech feature and avoid noise) from the at least one acoustic signal using the first transfer function to produce at least one denoised acoustic data stream (60, see col. 15 line 29-col.16 line 3 and col. 60 line 19-30).

Again, Holzrichter never discusses removing noise from acoustic signals. Holzrichter only discusses "removing 'post-glottal pressure induced vibrations' (i.e. the speech of the user) from the 'EM sensor signal' " (col. 24 lines 29-30).

> Consider claim 33 Holzrichter teaches a system for removing noise from acoustic signals, comprising at least one processor (see fig.3b (processing electronics)) coupled among at least one microphone (see fig.5 (52)) and at least one voicing sensor (43), wherein the at least one voicing sensor (43) detects human tissue vibration associated with voicing, wherein an absence of voiced information (unvoice) is detected during at least one period (such as, time frames)(see col.28 line 38-48) using the at least one voicing sensor, wherein at least one noise source signal is received during the at least one period using the at least one microphone (52), wherein the at least one processor generates at least one transfer function (57) representative of the at least one noise source signal, wherein the at least one microphone(52) receives at least one composite signal comprising acoustic and noise signals, and the at least one processor removes the noise signal(removing noise is inherent to speech recognition algorithm to extract the best speech feature and avoid noise) from the at least one composite signal using the at least one transfer function (57) to produce at least one denoised acoustic data stream(60, see col. 15 line 4-col.16 line 3 and col. 60 line 19-30).

Throughout Holzrichter's disclosure, Holzrichter fails to teach or suggest removing noise from acoustic signals. Holzrichter only discusses "removing 'post-glottal pressure induced vibrations' (i.e. the speech of the user) from the 'EM sensor signal' " (col. 24 lines 29-30). This illustrates not only that Holzrichter lacks claimed elements, but that Holzrichter is not attempting a teaching suggestive of the claims.

> Consider claim 35 Holzrichter teaches a signal processing system (see fig.3b, (processing electronic)) coupled among at least one user and at least one electronic device (see fig.3b, (processing electronic)), wherein the signal processing system (processing electronic) includes at least one denoising subsystem (see fig. 5) for removing noise from acoustic signals, the denoising subsystem (fig.5) comprising at least one processor coupled among at least one receiver and at least one sensor (43, EM sensor), wherein the at least one receiver is coupled to receive at least one acoustic signal; wherein least one sensor (43) detects human tissue vibration associated with human voicing activity (see col. 15 line 4-18), wherein the at least one processor generates a plurality of transfer functions (56, 57,59), wherein at least one first transfer function (56) representative of the at least one acoustic signal is generated in response to a determination that voicing information is absent (such as, unvoice) from the at least one acoustic signal for at least one specified period of time(such as, time frames)(see col.28 line 38-48), wherein noise is removed (removing noise is inherent to speech recognition algorithm to extract the best speech feature and avoid noise) from the at least one acoustic signal using the first transfer function to produce at least one denoised acoustic data stream (60, see col. 15 line 4-col.16 line 3 and col. 60 line 19-30).

Applicants respectfully reiterate that Holzrichter does not contain or suggest claimed elements in <u>any</u> of the claims. For example:

1. Holzrichter only uses a single microphone that is assumed to be noise free; while the claims include least two microphones that are assumed to be noisy;
2. Holzrichter does not calculate transfer functions between microphones (as in the claimed invention), but only Fourier transforms of clean acoustic data and "EM signals";
3. Holzrichter does not detect skin vibrations due to user speech; and
4. Holzrichter does not use a voice activity detection (VAD) signal to determine when to update the transfer functions between the microphones.

Rejections under 35 U.S.C. § 103

Claims 4, 19, and 32 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Holzrichter. Applicants respectfully traverse the rejection. The examiner states in paragraph 7 of the Office action:

> 7. Claims 4, 19 and 32 are rejected under 35 U.S.C. 103(a) as being unpatentable over Holzrichter (US PAT. 5,729,694).
>
> Consider claim 4 Holzrichter teaches implementing a microphone (acoustic sensor, e.g. microphone, col. 16 lines 14-15; col. 11 lines 29-30) coupled to a processor but does not teach implementing a plurality of microphones. At the time of the invention, it would have been obvious to one of ordinary skill in the art to implement a plurality of microphones for flexibility, thus gathering acoustic information in various areas instead of implementing one microphone in that is restricted to a certain area.
>
> Consider claims 19 and 32, they are essentially similar to claim 4 and are rejected for the reason stated above apropos to claim 4.

Applicants are uncertain of the meaning of "flexibility". The sentence goes on to state "thus gathering acoustic information in various areas instead of implementing one microphone in that it is restricted to a certain area". From the latter part of the sentence, Applicants can only conclude that "flexibility" has nothing to do with the use of two microphones as in the claimed system. The claimed system relies on two microphones to function. At least two microphones are the minimum number for implementing the claimed method and system. One of ordinary skill in the art would find no motivation to modify Holzrichter as suggested because Holzrichter does not teach or suggest the claimed nethod of processing data from ANY microphone (never mind more than one microphone) in the claimed manner or to achieve the claimed effect. The disclosure of Holzrichter is simply not enabling of a multiple-microphone method or system, and is not even suggestive of such a method or system. For all of these reasons, Applicants respectfully submit that claims 4, 19, and 32 would not have been obvious in view of Holzrichter.

Claim 31 was rejected under 35 U.S.C. § 103(a) as being unpatentable over Holzrichter in view of Sugiyama. Applicants respectfully traverse the rejection. The examiner states in paragraph 8 of the Office action that

> Holzrichter does not clearly teach the system of further comprising: dividing acoustic data of the at least one acoustic signal into a plurality of subbands; removing noise from each of the plurality of subbands using the at least one first transfer function, wherein a plurality of denoised acoustic data streams are generated; and 6 [sic] combining the plurality of denoised acoustic data streams to generate the at least one denoised acoustic data stream.

The Office action further states that Sugiyama teaches the system of further comprising: dividing (fig. 1 (50)); removing (6); and combining ((8) and col. 1, lines 12-35).

Applicants respectfully submit that the proposed combination does not result in claim 31. Sugiyama does not overcome the deficiencies of Holzrichter. For example, Sugiyama is not concerned with noise removal at all and so certainly fails to teach or suggest transfer functions. Sugiyama (5,517,435) simply implements a standard adaptive filter to do system identification, not noise removal. For all of these reasons, Applicants respectfully submit that the invention of claim 31 would not have been obvious is view of the cited references.

CONCLUSION

In view of the foregoing amendments and Remarks, Applicants respectfully submit that any objections and rejections have been overcome, and the claims are now allowable. Prompt allowance of the application is earnestly solicited. Examiner Lao is respectfully requested telephone the undersigned to facilitate resolution of any issues prior to allowance of the application.

AUTHORIZATION TO CHARGE DEPOSIT ACCOUNT

If there are any fees due and unpaid in this application, please charge our

Deposit Account No. 503616 for these fees.

Respectfully submitted,
Courtney Staniford & Gregory LLP

Date: July 23, 2007

Barbara B. Courtney, Reg. No. 42,442
Tel. 408-342-1902
Fax. 408-342-1909

Courtney Staniford & Gregory LLP
P.O. Box 9686
San Jose CA 95157